Achievement Testing Program

Provincial Report
June 1988 Administration

ATA LIBRARY

Alberta
EDUCATION
Student Evaluation and Records

September 1988

AE-CUR
ACH

DISTRIBUTION: Superintendents of Schools • School Principals and Teachers • The Alberta Teachers' Association • Alberta School Trustees' Association • Alberta Education • General Public upon Request

EXECUTIVE SUMMARY

DESCRIPTION AND ADMINISTRATION OF THE TESTS

The 1988 achievement tests were based on the prescribed objectives of the *Grade 3 Social Studies, Grade 6 English Language Arts,* and *Grade 9 Mathematics* provincial programs of studies.

The Grade 3 Social Studies Achievement Test consisted of 50 multiple-choice questions on knowledge and inquiry skills. The test was administered to 29 961 students.

The Grade 6 English Language Arts Achievement Test was a two-part test. Part A: Composition consisted of an assignment in which students were required to write a story based on a story starter. Part B: Reading consisted of 50 multiple-choice questions on a variety of literary and expository reading selections. In all, 32 069 students wrote the Grade 6 English Language Arts Achievement Test.

The Grade 9 Mathematics Achievement Test consisted of 75 multiple-choice questions on the six strands in the core program. The test was written by 25 290 students. A total of 1 867 students whose instruction was based on the new mathematics program wrote the pilot test.

As well, a Grade 6 French Language Arts Test was administered to 1 689 students and a Grade 9 Mathematics Test was translated into French and was administered to 773 students. These tests were available on an optional basis to serve those schools where the language of instruction is French. Only school and jurisdiction results were reported for these tests.

1988 ACHIEVEMENT TEST RESULTS

Grade 3 Social Studies

Student achievement in Grade 3 Social Studies in 1988 was about as high as expected. Results show that 80.0 per cent of the students who wrote the achievement test achieved an acceptable performance level and 18.8 per cent achieved an outstanding level. A comparison of 1984 results shows that achievement in 1988 was the same as in 1984.

~~#26019~~

AE-CUR
ACH

Grade 6 English Language Arts

In Grade 6 English Language Arts, student achievement was about as high as expected. Results show that 83.1 per cent of the students who wrote the test achieved an acceptable performance level and 9.9 per cent achieved at an outstanding level. A comparison with results in 1984 shows that, overall, achievement was higher in 1988.

Grade 9 Mathematics

Student achievement in Grade 9 Mathematics was higher than expected in 1988. Results show that 88.9 per cent of the students who wrote the achievement test achieved an acceptable level of performance and 24.3 per cent achieved at an outstanding level. A comparison with results in 1984 indicates that achievement was higher in 1988.

CONTENTS

PAGE

Section 1: Purpose of the Provincial Report 1

Section 2: The Student Achievement Testing Program 3

Section 3: Guidelines for Interpreting the Results of Achievement Tests 5

Section 4: Grade 3 Social Studies 9

Section 5: Grade 6 English Language Arts 17

Section 6: Grade 9 Mathematics 29

Section 7: Summary of Results 37

Section 8: Studies of Changes in Student Achievement Over Time 39

Appendix A: Standard-Setting Procedures 49

Appendix B: Equating Multiple-Choice Test Scores 50

Appendix C: Grade 6 English Language Arts: Written-Response Equating . . . 51

Achievement Testing Program: Provincial Report: Questionnaire . . . end page

---

LIST OF TABLES

TABLE

SECTION 4

PAGE

1 Grade 3 Social Studies: Achievement Test Blueprint 10

2 Grade 3 Social Studies: Students Tested, Absentees, and Exemptions 11

3 Grade 3 Social Studies: Raw Score Results By Reporting Category 12

4 Grade 3 Social Studies: Comparison of Results: Total and Partial Completion of the Test 13

5 Grade 3 Social Studies: Frequency Distribution of Raw Scores 14

6 Grade 3 Social Studies: Results for Individual Multiple-Choice Questions 15

7 Grade 3 Social Studies: Distribution of Jurisdiction Levels of Achievement 16

SECTION 5

8 Grade 6 English Language Arts: Achievement Test Blueprint Part A: Composition 18

9 Grade 6 English Language Arts: Achievement Test Blueprint Part B: Reading (Multiple Choice) 19

10 Grade 6 English Language Arts: Students Tested, Absentees, and Exemptions 20

11 Grade 6 English Language Arts: Frequency Distribution of Total Combined Scores 21

12 Grade 6 English Language Arts: Percentage Distribution of Scores/Part A: Composition 22

13 Grade 6 English Language Arts: Raw Score Results by Subtest Part B: Reading (Multiple Choice) 24

14 Grade 6 English Language Arts: Frequency Distribution of Raw Scores/Part B: Reading (Multiple Choice) 25

15 Grade 6 English Language Arts: Results for Individual Multiple-Choice Questions 26

16 Grade 6 English Language Arts: Distribution of Jurisdiction Levels of Achievement 27

TABLE PAGE

SECTION 6

17 Grade 9 Mathematics: Achievement Test Blueprint 30

18 Grade 9 Mathematics: Students Tested, Absentees, and Exemptions 31

19 Grade 9 Mathematics: Raw Score Results by Subtest. 32

20 Grade 9 Mathematics: Frequency Distribution of Raw Scores 33

21 Grade 9 Mathematics: Results for Individual Multiple-Choice Questions 34

22 Grade 9 Mathematics: Distribution of Jurisdiction Levels of Achievement 35

SECTION 7

23 Summary of Provincial Averages, Standard Deviations, and Performance: June 1988 Achievement Tests 37

SECTION 8

24 Grade 3 Social Studies: Multiple-Choice Comparison of the 1984 and the 1988 Test Results 41

25 Grade 3 Social Studies: Multiple-Choice Standard-Setting Results 41

26 Grade 6 English Language Arts: Multiple-Choice Comparison of the 1984 and the 1988 Test Results 43

27 Grade 6 English Language Arts: Multiple-Choice Standard-Setting Results 43

28 Grade 6 English Language Arts: Written-Response Frequencies of Scores of Satisfactory or Higher in Per Cent . 44

29 Grade 9 Mathematics: Multiple-Choice Comparison of the 1984 and the 1988 Test Results 46

30 Grade 9 Mathematics: Multiple-Choice Standard-Setting Results 46

SECTION 1:
PURPOSE OF THE PROVINCIAL REPORT

Following each administration of the achievement tests, a *Provincial Report* of the results is prepared. Superintendents and principals also receive reports of the test results for their students' achievement, which are parallel in format to the *Provincial Report*.

This *Provincial Report* is intended to provide information about the results achieved by students who wrote the June 1988 achievement tests. Specifically, these tests were Grade 3 Social Studies, Grade 6 English Language Arts, and Grade 9 Mathematics.

Provincial results of the Grade 6 French Language Arts Achievement Test and the French translations of the Grade 9 Mathematics Achievement Test and the Grade 9 Mathematics Pilot Test are NOT presented in this *Provincial Report*. Participation in the French Achievement Testing program is not mandatory. Since the population of students who wrote is self-selected, the results are not generalizable to the total population of Grade 6 French Language Arts students. Provincial statistics for the French Language Program, therefore, were not prepared. However, jurisdictions that chose to have their Francophone and/or French Immersion students write the Grade 6 French Language Arts Achievement Test or a French translation of the Grade 9 Mathematics Achievement Test will receive reports of jurisdiction and school results. Jurisdictions in which students wrote either the English or French version of the Grade 9 Mathematics Pilot Test will also receive reports of jurisdiction and school results. NOTE: While comparison of these results to provincial norms is interesting, such comparisons may not be valid.

This *Provincial Report* is intended to supplement the following Alberta Education documents:

- the school summary reports and the jurisdiction summary reports
- the *Student Achievement Testing Program Bulletins*, which were distributed in September 1987 and describe the achievement tests at each grade level
- the test specifications for each grade level provided by the Curriculum Design Branch and the Language Services Branch
- the achievement tests administered in 1988.

Information from all of these documents should enable teachers, principals, superintendents, and others to make judgments concerning student achievement at the provincial level and within schools and school jurisdictions.

Educators in each jurisdiction are encouraged to study the achievement test results carefully as an aid in determining the strengths and weaknesses of their instructional programs and resources. Once strengths and weaknesses have been determined, educators will be able to plot courses of action that will help to ensure that each student within a school or jurisdiction receives the best possible education.

It is strongly recommended that educators and others involved in interpreting results read "Section 3: Guidelines for Interpreting the Results of Achievement Tests" (see page 5) before interpreting the specific reports for their own jurisdiction or school.

Readers of this report are invited to complete the questionnaire at the end of this booklet and send it to the address indicated.

SECTION 2:
THE STUDENT ACHIEVEMENT TESTING PROGRAM

INTRODUCTION

The Student Achievement Testing Program is intended to provide Alberta Education, school jurisdictions, schools, and the public with information significant at the provincial and local levels about student knowledge, understanding, and skills in relation to program objectives. The Achievement Testing Program is NOT intended to provide information to be used in connection with student placement or promotion, nor as the basis for teacher evaluation.

The achievement tests are administered on a four-year cycle in four subject areas: language arts, social studies, mathematics, and science; and at three grade levels: 3, 6, and 9.

The achievement tests are specific to the programs of studies prescribed by the Minister of Education. Curriculum specifications for each subject area, prepared by the Curriculum Design Branch and the Language Services Branch of Alberta Education, identify the major content areas, the specific learning objectives within each content area, and the emphasis that each objective is to receive. The questions on each achievement test reflect the emphases outlined in these curriculum specifications.

Classroom teachers from across the province are extensively involved in developing and field testing the questions. The student responses are analysed after field testing to determine each question's discriminating power and level of difficulty. Questions may undergo several revisions before they appear on an achievement test.

The final draft of each test is examined by an Achievement Test Review Committee that includes representatives of The Alberta Teachers' Association, the Conference of Alberta School Superintendents, Alberta's post-secondary institutions, and Alberta Education.

For a more complete description of the achievement test development process, please refer to the Student Evaluation and Records Branch brochure entitled *Developing Achievement Tests*. Copies of this brochure can be obtained from the Student Evaluation and Records Branch.

EXEMPTIONS FROM THE ACHIEVEMENT TESTING PROGRAM

All students who were taught a subject at the grade level being tested were expected to participate in the Achievement Testing Program.

Only those students for whom participation in the Achievement Testing Program was inappropriate should have been excused from writing. These are students who

- were enrolled in an approved program that has been designed for special needs students
- were taught the subject being tested in another semester or year
- were enrolled in an English as a Second Language program
- were being taught the specific subject being tested (mathematics, language arts, or social studies) in a language other than English or French.

Any exemptions were identified by the principal and approved by the superintendent of schools. Exemptions for any reasons other than those outlined above were approved by the Director of the Student Evaluation and Records Branch.

SECTION 3: GUIDELINES FOR INTERPRETING THE RESULTS OF ACHIEVEMENT TESTS

JURISDICTION AND SCHOOL REPORTS, AND INDIVIDUAL STUDENT SUBTEST RESULTS

In addition to their use in monitoring student achievement for the province as a whole, the results of the achievement tests are useful for comparing the achievement of students in a particular jurisdiction or school with the achievement of students in the entire province of Alberta. Superintendents and principals receive confidential reports of results achieved by the students in their particular jurisdiction or school. These are

- the *Jurisdiction Summary Report,* which contains *jurisdiction* equivalents of the provincial results that are given in all tables in this report
- the *School Summary Report,* which contains *school* equivalents of the provincial results that are given in all tables in this report.

In addition to these reports, each superintendent and principal receives a report entitled *Individual Student Subtest Results.* This report lists the results achieved by each student in the jurisdiction or school on the subtests or reporting categories of the achievement tests.

Included in each jurisdiction and school report are jurisdiction or school average scores for each reporting category or subtest. Each of these scores may be compared to the provincial average for the same reporting category or subtest to determine if differences exist.

The importance of differences that may exist between jurisdiction or school averages and provincial averages is not always clear, however.

To aid in the interpretation of differences between the averages, jurisdiction and school reports indicate when a difference is unlikely to be due to chance variation. For the purposes of the Achievement Testing Program, the 95% confidence interval is used. That is, if the probability is less than one in 20 that a difference is due to chance, this difference is very likely a real difference, and the jurisdiction or school average is classified as being different from (significantly above or significantly below) the provincial average.

Because achievement levels are calculated by taking jurisdiction or school size into consideration, two jurisdictions or schools with the same averages but of different sizes may be classified differently. The larger jurisdiction or school would be more likely to be above or below average because the amount of chance variation would be less in larger jurisdictions or schools, and the actual difference would represent a larger variation from the provincial average.

USE OF THE REPORTS

The reports may be used as part of the information for

- evaluating educational programs in each subject at each grade level
- improving the quality of educational programs
- identifying strengths and weaknesses of school and jurisdiction programs by comparing their results with provincial results.

The reports are NOT intended to be used as the basis for

- evaluating teacher performance
- comparing performance between or among schools.

Administrators in each jurisdiction should apply separate locally developed teacher, school, and school system evaluation policies to the tasks of evaluating teacher and school performance.

The information provided in the reports is factual regarding *what* has happened as a result of the administration of the tests. The interpretation of this information -- hypothesizing *why* results are as they are -- involves many complex considerations of the factors and variables that contribute to achievement.

In addition, it must be noted that the information in this report is itself limited to selected objectives of the *Program of Studies*. Many important aspects of learning cannot be measured by the achievement tests, which are time-limited paper and pencil tests.

FACTORS THAT MAY AFFECT STUDENT ACHIEVEMENT

There are many factors or variables that may affect student achievement:

1. Environment
    - community environment
    - school environment
    - socioeconomic background
    - family circumstances
2. Student Factors
    - ability
    - attitude
    - motivation
    - aspiration
    - academic background
    - learning style
3. Resources (availability and appropriateness)
    - programs of study
    - curriculum guides
    - resource materials
    - library services
    - current textbooks
    - references
4. Instruction
    - teacher qualifications
    - teacher experience
    - professional development
    - teacher morale
    - teaching strategies
    - hours of instruction
    - staff turnover
    - amount of homework assigned
    - communication of teacher expectations

A SYSTEMATIC APPROACH FOR THE EFFECTIVE USE OF ACHIEVEMENT TEST RESULTS

Achievement test results can be used constructively as one means of improving the quality of education. Some steps for a systematic use of achievement test results are listed below:

1. Compare test results for a jurisdiction or school to the provincial results. Be sure that your comparisons include the
    - total test score
    - total and subtest scores for multiple-choice questions
    - total and subtest scores for written-response assignments (when appropriate)
    - individual multiple-choice question results
    - individual written-response assignment results (when appropriate).
2. Note any patterns, anomalies, and/or inter-relationships in the results.
3. Hypothesize relationships between your observations and factors (see page 6) that may have had an effect on achievement.
4. Consider and plan an action that will help improve the quality of education for students.

AN ADMINISTRATIVE MODEL FOR THE EFFECTIVE USE OF ACHIEVEMENT TEST RESULTS

The following is *one* model that may be useful for those wishing to develop a constructive system for interpreting achievement test results. This model is based on work done by Medicine Hat School District #76.

Basic Principles

1. It is desirable and feasible for teachers and school administrators to make use of achievement test results in analysing the performance of their own students.
2. It is more constructive for schools to develop their own analyses, interpretations, and action plans than to have these imposed externally.
3. The impact of factors such as those listed on page 6 should be analysed and discussed when reviewing achievement test results.
4. Subtest or reporting category results are usually more informative than total test scores.
5. Generalizations should be stated with caution and should be supported by evidence that is independent of achievement test results.
6. It is neither desirable nor productive to compare the results of different schools.
7. Achievement tests measure many of the objectives specified by the curriculum. *However, skills and concepts that are not measured by the achievement tests are also to be taught and evaluated at the local level.*
8. Staff discussions as well as written reports are useful means of ensuring that results are appropriately interpreted and used.

Suggested Content for Interpretations of Individual School Results

1. Subject, grade level, and date of achievement test administration
2. Number of students who wrote the achievement test
3. Profiles of students or groups who wrote the achievement test, which include
   - previous performance
   - number of students repeating the grade, etc.
4. Program emphases, such as hours of instruction, skill and content emphases
5. Instructional practices, such as methodology, resources, and the relationship between the program offered and the provincial curriculum
6. Program objectives not measured by the achievement test
7. School results compared to provincial results on subtests
8. Current school results compared to those of previous administrations
9. Discussion of actual and expected subtest results, with recommended action
10. Discussion of item results, identification of common student errors, and suggestions of ways for reducing the misunderstanding that leads to these errors
11. Recommendations for the following year or semester
12. Summary and concluding comments

Suggested Reporting Structure

1. Teachers and/or the principal analyse the results and prepare a written report about each administration of an achievement test.
2. The principal reviews and signs the report.
3. The report is shared with central office supervisory personnel.
4. The appropriate central office supervisory personnel prepare a written response to the report, with copies of the response going to the teachers and principal.
5. If possible, all involved staff meet to discuss the report and the response.
6. Reports are used to improve the program and maximize future opportunities for student success.
7. When large differences exist between expected and actual achievement test results over time, consideration should be given to conducting a formal program evaluation.

SECTION 4:
GRADE 3 SOCIAL STUDIES

THE GRADE 3 SOCIAL STUDIES ACHIEVEMENT TEST

General Description

The Grade 3 Social Studies Achievement Test consisted of 50 multiple-choice questions that were answered directly in the test booklet in machine-scorable format. The time allotted for writing the test was 60 minutes.

The test was divided into three equal portions. Rest breaks of approximately 10 minutes were given to students between each 20-minute portion of the test.

Content of the Test

The Grade 3 Social Studies Achievement Test was based on the *1981 Alberta Social Studies Curriculum*. All test questions were drawn from the content of the three topics prescribed for Grade 3. These topics are

- Topic A: Interdependence of Communities in Canada Today
- Topic B: Lifestyles of Canadians in Other Times
- Topic C: Lifestyles in Culturally Distinctive Communities.

Topic A contains 17 questions; Topic B, 17 questions; and Topic C, 16 questions.

Content emphases were drawn from the *Grade 3 Social Studies Curriculum Specifications*.

The Grade 3 Social Studies Achievement Test measured value, knowledge, and skill objectives. Objectives related to the development of attitudes and participation skills did not form part of this test. The weighting allocated to the development of these objectives in the curriculum specifications was reassigned to the remaining objectives on a prorated basis.*

Table 1 presents the blueprint used to develop the Grade 3 Social Studies Achievement Test. Classification by reporting category for each question included in the test is indicated in the table.

---

*Although these objectives are important, questions based on such objectives yield different data than are ordinarily associated with student performance on an achievement test.

Table 1
Grade 3 Social Studies
Achievement Test Blueprint

	Concept Reporting Category				
Process Reporting Category	Value concepts: Competing values and value positions	Interdependence of Communities in Canada Today: Facts, concepts, and generalizations related to the interdependence of communities in Canada today	Lifestyles of Canadians in Other Times: Facts, concepts, and generalizations related to the lifestyle of Canadians in other times	Lifestyles in Culturally Distinctive Communities: Facts, concepts, and generalizations related to lifestyles in culturally distinctive communities	Total(%)
Recall and Comprehension: The student will be able to recall or recognize data and transform data into other words.	3,20,21, 44,48	6,7,8,9, 10,11,12	25,26,29,30, 31,32,33,34	35,36,39,46, 47,49,50	54
Inquiry Skills I: The student will be able to use skills in the early stages of the Inquiry Process, such as identifying the issue and gathering and organizing data.		1,2,4,14, 17	18,22,23,28	37,38,40,41	26
Inquiry Skills II: The student will be able to use skills in the later stages of the Inquiry Process, such as analysing data and resolving the issue.		13,15,16	19,24,27	42,43,45	18
Valuing Skills: The student will be able to analyse competing values.	5				2
Total(%)	12	30	30	28	100

ABSENTEES AND EXEMPTIONS

Table 2 presents the number and percentage distribution of students who wrote the Grade 3 Social Studies Achievement Test, as well as the number and percentage distribution of students who were absent or who were exempted from writing for various reasons.

Table 2
Grade 3 Social Studies
Students Tested, Absentees, and Exemptions

Category	Number of Students	Per Cent
Students Tested	29 961	88.2
Students Absent	968	2.9
Students Exempt:		
• Special Needs	686	2.0
• Not Taught Subject This Term	378	1.1
• ESL Classes	315	0.9
• Language of Instruction NOT English	659	1.9
• Other (Approved by Director)	0	0.0
Total Exempt	2 038	5.9

Note: The data are incomplete as not all schools returned absentee and exemption reports. As well, 1 008 students wrote only one or two of the three topics. The provincial report does not include the results for these students.

RESULTS

Information in all tables contained in this section is based on the results achieved by 29 961 students. Since over 29 400 of these students completed all questions on the test, it was concluded that sufficient time was allotted for writing the test.

The minimum acceptable and outstanding performance levels of achievement were established by a standard-setting procedure (see Appendix A). For the Grade 3 Social Studies Achievement Test, the standards established are such that to meet

- the minimum acceptable performance level, students must achieve a raw score of 26 out of 50
- the outstanding performance level, students must achieve a raw score of 43 out of 50.

Based on these standards, the results revealed that

- 80.0 per cent of students performed at or above the minimum acceptable performance level
- 18.8 per cent of students performed at or above the outstanding performance level.

These levels of performance were about as high as expected.

Provincial scores, in percentage points, were as follows:

- Provincial Average - 67.6
- Standard Deviation - 18.2

The KR-20 reliability coefficient for the Grade 3 Social Studies Achievement Test was 0.89.

Provincial averages for these reporting categories and for the total test are presented in Table 3. As outlined in the blueprint presented on page 10, the questions on the Grade 3 Social Studies Achievement Test have been grouped according to reporting categories. Provincial averages were computed and rounded to one decimal. Consequently, the sum of the averages for the reporting categories is not exactly the same as the average for the total test.

Table 3
Grade 3 Social Studies
Raw Score Results by Reporting Category

Reporting Category	Number of Questions	Average	Standard Deviation
Total Test	50	33.8	9.1
Topic A: Interdependence of Communities in Canada Today	17	11.7	3.3
Topic B: Lifestyles of Canadians in Other Times	17	11.4	3.2
Topic C: Lifestyles in Culturally Distinctive Communities	16	10.7	3.8
Recall and Comprehension			
All Topics	22	14.8	4.4
Topic A	7	4.8	1.8
Topic B	8	5.3	1.8
Topic C	7	4.7	1.9
Value Concepts and Valuing Skills (all topics)	6	4.0	1.5
Inquiry Skills I (all topics)	13	8.7	2.5
Inquiry Skills II (all topics)	9	6.3	2.1

Although levels of performance in the different reporting categories appeared to show some variation, caution is advised in comparing them. The sets of questions that make up each category were not selected to be equal in average level of difficulty; therefore, differences may have been due to variations in question difficulty rather than in student performance. The provincial averages can be used in combination with jurisdiction and school results to detect patterns of relative strength or weakness in achievement in each of the reporting categories.

Results for Students Completing Only Portions of the Test

Table 4 presents and compares by topic the results for students who wrote only portions of the test with the results for students who completed the total program and therefore wrote the entire test. Students in some split-grade classes did not receive instruction in all three topics. These students were exempt from writing those topics on the test that they had not been taught. Partially exempt student scores are not included in provincial, jurisdiction, and school statistics.

Table 4
Grade 3 Social Studies
Comparison of Results
Total and Partial Completion of the Test

	Students Completing the Total Test			Students Completing Only Portions of the Total Test		
	No. of Students	Mean	Standard Deviation	No. of Students	Mean	Standard Deviation
Topic A	29 961	11.5	3.6	510	12.0	3.2
Topic B	29 961	11.3	3.4	687	11.4	3.2
Topic C	29 961	10.6	3.9	680	11.5	3.4

There was only a slight difference in the results among the groups on all three topics.

The percentage of students who obtained each score on the Grade 3 Social Studies Achievement Test (relative frequency) and the percentage of students scoring at or below each score (cumulative frequency) are presented in Table 5.

Table 5
Grade 3 Social Studies
Frequency Distribution of Raw Scores

Raw Score	Relative Frequency (%)	Cumulative Frequency (%)
0	0.0	0.0
1	0.0	0.0
2	0.0	0.0
3	0.0	0.0
4	0.0	0.0
5	0.0	0.1
6	0.0	0.1
7	0.0	0.1
8	0.1	0.2
9	0.1	0.3
10	0.2	0.5
11	0.3	0.8
12	0.4	1.2
13	0.6	1.8
14	0.6	2.4
15	0.8	3.2
16	1.0	4.2
17	1.3	5.5
18	1.3	6.8
19	1.3	8.1
20	1.5	9.6
21	1.8	11.5
22	2.0	13.4
23	2.0	15.4
24	2.3	17.7
25	2.3	20.0
26	2.4	22.4
27	2.8	25.2
28	2.8	27.9
29	3.0	30.9
30	3.0	33.9
31	3.2	37.1
32	3.4	40.5
33	3.8	44.3
34	3.7	48.0
35	3.8	51.8
36	4.1	55.9
37	4.0	59.9
38	4.2	64.1
39	4.3	68.5
40	4.3	72.7
41	4.3	77.0
42	4.2	81.2
43	3.9	85.1
44	3.8	89.0
45	3.3	92.3
46	2.9	95.2
47	2.2	97.5
48	1.5	99.0
49	0.8	99.8
50	0.2	100.0

Table 6 presents the percentage of students choosing each alternative (A, B, C, and D) for each multiple-choice question on the Grade 3 Social Studies Achievement Test. The correct response (key) for each question is also identified. The results presented in Table 6 can best be used with results presented in jurisdiction and school reports to interpret strengths and weaknesses of local programs.

Table 6
Grade 3 Social Studies
Results for Individual Multiple-Choice Questions

		Distribution of Responses* (%)						Distribution of Responses* (%)			
Item	Key	A	B	C	D	Item	Key	A	B	C	D
1	A	87.3	5.0	3.8	3.0	26	D	12.1	4.9	5.4	74.6
2	C	10.5	28.0	53.8	7.2	27	B	3.8	81.9	3.7	8.5
3	D	15.3	4.1	9.4	70.5	28	A	48.8	3.4	25.8	19.0
4	B	24.5	59.5	9.8	4.3	29	B	7.7	59.9	18.0	12.1
5	C	5.8	20.1	56.8	16.8	30	C	2.2	4.9	84.9	6.5
6	B	4.7	77.8	6.9	9.7	31	A	50.0	3.7	7.4	37.1
7	D	9.1	23.4	9.3	56.8	32	D	6.0	10.0	10.8	70.6
8	C	17.0	3.5	71.7	6.5	33	D	12.7	5.2	8.4	71.7
9	C	4.2	17.8	69.8	5.9	34	B	7.0	63.4	12.2	14.8
10	D	11.7	12.6	8.8	65.2	35	C	4.3	25.0	60.1	9.9
11	A	70.7	2.6	5.5	18.3	36	B	6.5	71.8	8.5	12.8
12	B	14.5	68.9	11.4	4.1	37	A	90.1	5.8	0.9	2.7
13	A	78.1	3.6	12.6	4.3	38	D	7.8	14.7	23.9	53.0
14	C	10.9	21.1	53.5	12.7	39	D	10.1	6.0	10.0	72.9
15	D	11.1	7.1	8.1	70.8	40	D	8.9	4.8	17.9	67.7
16	A	82.9	5.9	5.4	4.2	41	A	70.2	14.1	10.9	3.8
17	B	2.6	73.3	11.6	5.4	42	B	12.1	56.0	11.6	18.8
18	A	71.1	15.2	4.0	2.1	43	A	65.8	14.0	7.5	10.5
19	D	6.2	4.2	18.1	70.9	44	B	10.8	69.6	11.1	6.7
20	D	6.7	4.6	6.0	81.9	45	D	7.8	7.2	16.9	65.7
21	A	68.7	3.6	17.9	9.3	46	D	8.0	6.8	22.3	61.6
22	C	7.3	5.3	75.5	11.3	47	D	10.4	12.5	8.7	66.4
23	B	9.6	62.2	10.5	16.9	48	A	57.3	4.9	5.4	30.3
24	D	4.3	10.9	27.6	55.8	49	B	4.8	79.7	10.4	3.7
25	C	5.9	36.5	52.9	4.1	50	C	15.0	13.4	59.5	10.6

*The sum of the percentages for each question may be less than 100% because the No Response category is not included. The No Response category does not exceed 7.5% for any one of these questions.

Questions on the Grade 3 Social Studies Achievement Test ranged from those that students found to be difficult, such as question 28 with only 48.8% of the students answering correctly, to those that were quite easy for students, such as question 37 with 90.1% of the students answering correctly. An example of a question of average difficulty is question 40 with 67.7% of the students answering correctly.

Table 7 indicates the percentage of jurisdictions classified as significantly below, not different from, and significantly above the provincial average (see page 5) for each reporting category and for the total test.

Table 7
Grade 3 Social Studies
Distribution of Jurisdiction Levels of Achievement

Reporting Category	Distribution of Jurisdiction Levels of Achievement		
	Below Provincial Average(%)	Not Different from Provincial Average(%)	Above Provincial Average(%)
Total Test	22.6	51.3	26.1
Topic A: Interdependence of Communities in Canada Today	22.2	52.6	25.2
Topic B: Lifestyles of Canadians in Other Times	16.5	59.1	24.3
Topic C: Lifestyles in Culturally Distinctive Communities	27.4	51.7	20.9
Recall and Comprehension			
All Topics	24.8	50.4	24.8
Topic A	22.2	54.3	23.5
Topic B	17.8	62.6	19.6
Topic C	23.9	53.9	22.2
Value Concepts and Valuing Skills (all topics)	16.1	57.0	27.0
Inquiry Skills I (all topics)	18.7	59.1	22.2
Inquiry Skills II (all topics)	19.1	62.6	18.3

THE GRADE 6 ENGLISH LANGUAGE ARTS ACHIEVEMENT TEST

General Description

The Grade 6 English Language Arts Achievement Test was a two-part test. Part A: Composition was an 80-minute writing test consisting of a story starter and instructions for the student to finish writing the story. Part B: Reading (Multiple Choice) was a 60-minute reading test consisting of 50 multiple-choice questions based on reading selections taken from fiction, nonfiction, and poetry.

The Grade 6 English Language Arts Achievement Test was designed to reflect the Grade 6 Language Arts curriculum specifications that have been developed from the *Program of Studies for Elementary School 1978* (amended 1982). The scope of the Grade 6 English Language Arts Achievement Test was limited to the writing and the reading components of the program. The components of viewing, listening, and speaking were reflected in teachers' evaluations of students.

Content of Part A: Composition

The Part A: Composition booklet included pages that were labelled IDEAS/ PLANNING, PLANNING AND DRAFTING, and REVISED WORK. This format was designed to reflect the writing process model. Although specific marks for planning and drafting were not allotted, markers were advised to take planning and drafting into consideration when scoring student work.

Compositions were scored for Content, Development, Sentence Structure, Vocabulary, and Conventions.

Table 8 on page 18 presents the blueprint used in the development of Part A: Composition. This table outlines the reporting categories (scoring scales used in scoring the test), a description of the writing assignment, and an indication of the range of marks.

Content of Part B: Reading (Multiple Choice)

Whenever possible, reading selections chosen to be included in Part B: Reading (Multiple Choice) were complete works containing a beginning, a middle, and an end. Reading selections were chosen to reflect the interests of the majority of Grade 6 students and to be of appropriate difficulty for Grade 6 students. Extensive use was made of Canadian content.

Questions were developed to test how well the students understood and analysed the reading selections, and made judgments about form and content. Only questions dealing with significant aspects of the reading selections were used.

Questions were classified according to three cognitive levels: Literal Understanding (9 questions), Inferential Understanding (33 questions), and Evaluation (8 questions). By considering cognitive level when developing a test, the Student

Evaluation and Records Branch ensures that students will use a variety of mental activities as they write the test. Questions listed under Literal Understanding should use skills of recall and recognition; those listed under Inferential Understanding should elicit skills of analysis, interpretation, and extrapolation; and questions listed under Evaluation should draw forth judgmental skills.

Table 9 on page 19 presents the blueprint used to develop Part B: Reading (Multiple Choice). Classification by reporting category and process category for each question included in Part B: Reading (Multiple Choice) is indicated in the table.

Table 8
Grade 6 English Language Arts
Achievement Test Blueprint
Part A: Composition

Reporting Category (Scoring Guide)	Description of Writing Assignment	Range of Marks
Content (Selecting details to achieve a purpose) The student will be able to select details to describe characters and settings that are appropriate within the context or terms of reference established by the student.	The writing assignment follows a story starter that is to be read by each student. The assignment sets a specific task but allows the student to be imaginative in selecting details, actions, and events.	5 - Excellent 4 - Proficient 3 - Satisfactory 2 - Limited 1 - Poor INS - Insufficient/ No Response
Development (Organizing details into a coherent whole) The student will be able to place events in a coherent sequence.		
Sentence Structure (Structuring sentences effectively) The student will be able to use a variety of sentence structures effectively in writing.		
Vocabulary (Selecting and using words and expressions correctly and effectively) The student will be able to use words and expressions effectively in writing.		
Conventions (Using the conventions of language correctly and effectively) The student will be able to communicate clearly in writing by adhering to appropriate spelling, grammar, punctuation, and capitalization.		

Table 9
Grade 6 English Language Arts
Achievement Test Blueprint
Part B: Reading (Multiple Choice)

Reporting Category	Curriculum Specifications Reference	Process Category			Totals(%)
		Literal Understanding	Inferential Understanding	Evaluation	
Identifying the Organization of Ideas					
The student will be able to attend to and analyse elements of the author's organization, such as sequence, purpose, comparison, cause/effect, and imagery in a reading selection.	I.,A.,9,13,14 II.,A.,5,8,10 II.,B.,3 III.,B.,1 IV.,A.,1,2,3 V.,A.,2	21	12,13, 22,24, 44,48	19	16
Analysing Details					
The student will be able to attend to and analyse the interrelationship of the details in a reading selection.	I.,A.,2,9,11 II.,A.,2,7 II.,B.,3	1,2,9, 15,23, 31,33	3,5,8, 11,14, 27,28, 30,37, 38,42, 46	10,17, 41	44
Associating Meanings					
The student will be able to associate meanings of words and expressions in context and evaluate appropriateness of word choice.	I.,A.,3 II.,A.,3 II.,B.,2 IV.,B.,1	32	4,7,16, 20,35, 39,43, 45,47	25,36	24
Synthesizing Ideas					
The student will be able to synthesize the information within the reading selection in order to construct meaning. The student should be able to synthesize ideas from the entire reading selection in order to deduce the main idea and to predict plausible outcomes or conclusions.	I.,A.,1 II.,A.,1,7 II.,B.,3 II.,C.,1 II.,D.,1,2 IV.,A.,1,2,3		6,18, 26,29, 34,40	49,50	16
Totals(%)		18	66	16	100

ABSENTEES AND EXEMPTIONS

Table 10 presents the number and percentage distribution of students who wrote the Grade 6 English Language Arts Achievement Test, as well as the number and percentage distribution of students who were absent or who were exempted from writing for various reasons.

Table 10
Grade 6 English Language Arts
Students Tested, Absentees, and Exemptions

Category	Number of Students	Per Cent
Students Tested	32 069	94.2
Students Absent	993	2.9
Students Exempt:		
• Special Needs	695	2.0
• Not Taught Subject This Term	17	0.1
• ESL Classes	219	0.6
• Language of Instruction Not English	0	0.0
• Other (Approved by Director)	18	0.1
Total Exempt	949	2.8

Note: The data are incomplete as not all schools returned absentee and exemption reports.

RESULTS FOR THE TOTAL TEST

Results show that 83.1 per cent of the students who wrote the test achieved an acceptable performance level and 9.9 per cent achieved an outstanding level. It was concluded that student achievement was about as high as expected. (Please refer to pages 37 and 38 as well.)

To calculate a total test score for students, Part A: Composition and Part B: Reading (Multiple Choice) were given equal weightings. A summary score for Part A: Composition was calculated by summing up the scores for each of the five written-response scales, thus giving each scale equal weighting in the summary score. The average total score for the test was 62.5%.

Table 11 on the next page shows the frequency distribution of total combined scores. The scores are expressed in per cent.

Table 11
Grade 6 English Language Arts
Frequency Distribution of Total Combined Scores

Raw Score	Relative Frequency (%)	Cumulative Frequency (%)	Raw Score	Relative Frequency (%)	Cumulative Frequency (%)
0	0.0	0.0	51	1.9	24.1
1	0.0	0.0	52	2.0	26.2
2	0.0	0.0	53	2.0	28.2
3	0.0	0.0	54	2.0	30.2
4	0.0	0.0	55	2.1	32.3
5	0.0	0.0	56	2.2	34.5
6	0.0	0.0	57	2.2	36.8
7	0.0	0.0	58	2.3	39.0
8	0.0	0.0	59	2.3	41.4
9	0.0	0.0	60	2.4	43.8
10	0.0	0.0	61	2.6	46.4
11	0.0	0.0	62	2.5	48.9
12	0.0	0.1	63	2.5	51.4
13	0.0	0.1	64	2.5	53.9
14	0.0	0.1	65	2.6	56.5
15	0.0	0.1	66	2.4	58.9
16	0.0	0.2	67	2.5	61.4
17	0.0	0.2	68	2.5	63.9
18	0.0	0.2	69	2.3	66.2
19	0.0	0.3	70	2.2	68.5
20	0.1	0.3	71	2.3	70.8
21	0.1	0.4	72	2.3	73.1
22	0.1	0.5	73	2.2	75.3
23	0.1	0.6	74	2.1	77.4
24	0.1	0.8	75	2.0	79.4
25	0.2	0.9	76	1.8	81.2
26	0.1	1.1	77	1.7	82.9
27	0.1	1.4	78	1.7	84.5
28	0.2	1.4	79	1.5	86.1
29	0.2	1.7	80	1.5	87.6
30	0.3	1.9	81	1.3	88.9
31	0.3	2.3	82	1.2	90.1
32	0.4	2.6	83	1.1	91.2
33	0.4	3.0	84	1.2	92.4
34	0.5	3.5	85	1.0	93.2
35	0.5	4.0	86	0.9	94.3
36	0.6	4.5	87	0.8	95.1
37	0.7	5.2	88	0.7	95.8
38	0.8	6.1	89	0.7	96.5
39	0.9	6.9	90	0.6	97.1
40	1.0	7.9	91	0.6	97.7
41	1.0	8.9	92	0.5	98.1
42	1.1	10.0	93	0.4	98.6
43	1.2	11.2	94	0.4	99.0
44	1.4	12.6	95	0.3	99.3
45	1.4	14.0	96	0.3	99.6
46	1.4	15.4	97	0.2	99.8
47	1.5	16.9	98	0.0	99.9
48	1.7	18.6	99	0.0	100.9
49	1.7	20.3	100	0.0	100.0
50	1.8	22.2			

RESULTS FOR PART A: COMPOSITION

The average summary written-response score for Part A: Composition was 15.1 possible marks out of 25. This score was calculated by summing up the scores on each of the five five-point written-response scales.

Table 12 presents the percentage distribution of scores (scale points) on Part A: Composition.

Results for Part A: Composition are most clearly understood in the context of the assignment responded to by students and in the context of the scoring descriptors. Complete scoring guides are available in the *Student Achievement Testing Program Bulletin: Grade 6 English Language Arts 1987-88 School Year.*

All schools should also have extra copies of Part A: Composition to be used in conjunction with information provided in Table 12.

The results outlined in Table 12 are perhaps best considered in terms of the percentage of students that markers have judged to have presented work that was "Satisfactory" or better for any reporting category.

Please note that students do better on some dimensions of the task than on others. It is possible to draw conclusions about local program strengths and weaknesses by comparing local percentages of "Satisfactory" or better scores on each reporting category to the provincial averages.

Table 12
Grade 6 English Language Arts
Percentage Distribution of Scores
Part A: Composition

Score (Scale Points)	Reporting Category				
	Content	Development	Sentence Structure	Vocabulary	Conventions
5 (Excellent)	7.5	6.1	6.8	6.4	8.1
4 (Proficient)	21.7	20.6	21.7	18.5	23.6
3 (Satisfactory)	40.4	39.0	48.4	52.5	41.5
2 (Limited)	23.4	28.1	18.3	18.9	21.0
1 (Poor)	6.3	5.6	4.2	3.0	5.1
INS (Insufficient/ No Response)	0.6	0.6	0.6	0.6	0.6

In responding to the assignment, 69.6% of the students scored "Satisfactory" or better for Content. Similar results for the other reporting categories of the assignment are 65.7% for Development, 76.9% for Sentence Structure, 77.4% for Vocabulary, and 73.2% for Conventions. Relatively low percentages of students achieving satisfactory or better on the various strands are the result of the changes made to 1988 scoring guides and the nature of the assignment. This is especially true of the results for Content. Through standard-setting procedures and a special comparison study of expectations for achievement set in 1984, it is clear that students have achieved at least as well, and possibly better, in 1988. (Please refer to pages 44 and 45 as well.)

Although the papers were scored on a one-marker system, 200 papers were recirculated so that a second set of scores was available for these papers to confirm scoring consistency. Of the scores awarded on the second reading, 88.5% were identical to the original score on the same scale or varied by only one point. It is important to note that the one-marker system produces results that are reliable for groups of 25 or more students. Achievement test scores are not intended to be reliable for individual students.

RESULTS FOR PART B: READING (MULTIPLE CHOICE)

Information in all tables contained in this section is based on the results achieved by 32 069 students. Since over 31 600 of these students completed Part B: Reading (Multiple Choice), it was concluded that sufficient time was allotted for writing the test.

The minimum acceptable and outstanding performance levels of achievement were established by a standard-setting procedure (see Appendix A). For Part B: Reading (Multiple Choice), the standards established are such that to meet

- the minimum acceptable performance level, students must achieve a raw score of 23 out of 50
- the outstanding performance level, students must achieve a raw score of 43 out of 50.

Based on these standards, the results revealed that

- 83.3 per cent of students performed at or above the minimum acceptable performance level
- 14.9 per cent of students performed at or above the outstanding performance level.

These levels of performance were about as high as expected.

Provincial scores, in percentage points, were as follows:

- Provincial Average - 64.7
- Standard Deviation - 18.3

The KR-20 reliability coefficient for Part B: Reading (Multiple Choice) was 0.89.

As outlined in the blueprint presented on page 19, the questions on Part B: Reading (Multiple Choice) have been grouped according to subtests, reporting categories, and process categories.

Provincial averages for each of these subtests and for Part B as a whole are presented in Table 13. Provincial averages were computed and rounded to one decimal.

Table 13
Grade 6 English Language Arts
Raw Score Results by Subtest
Part B: Reading (Multiple Choice)

Subtest	Number of Questions	Average	Standard Deviation
Total Part B	50	32.4	9.2
Reporting Category			
Identifying the Organization of Ideas	8	4.7	1.9
Analysing Details	22	15.0	4.2
Associating Meanings	12	7.6	2.7
Synthesizing Ideas	8	5.1	1.7
Process Category			
Literal Understanding	9	6.7	1.9
Inferential Understanding	33	21.0	6.4
Evaluation	8	4.7	1.8

Although levels of performance in the different reporting categories and process categories appeared to show some variation, caution is advised in comparing them. The sets of questions that make up each category were not selected to be equal in average level of difficulty; therefore, differences may have been due to variations in question difficulty rather than in student performance. The provincial averages can be used, however, in combination with jurisdiction and school results to detect patterns of relative strength or weakness in achievement in each of the categories.

The percentage of students who obtained each score on Part B: Reading (Multiple Choice) (relative frequency) and the percentage of students scoring at or below each score (cumulative frequency) are presented in Table 14.

Table 14
Grade 6 English Language Arts
Frequency Distribution of Raw Scores
Part B: Reading (Multiple Choice)

Raw Score	Relative Frequency (%)	Cumulative Frequency (%)
0	0.0	0.0
1	0.0	0.0
2	0.0	0.0
3	0.0	0.0
4	0.0	0.0
5	0.0	0.0
6	0.0	0.0
7	0.0	0.1
8	0.1	0.2
9	0.2	0.3
10	0.3	0.6
11	0.4	1.0
12	0.6	1.6
13	0.7	2.3
14	0.8	3.1
15	1.1	4.2
16	1.2	5.4
17	1.3	6.8
18	1.6	8.3
19	1.8	10.1
20	1.9	12.0
21	2.2	14.2
22	2.4	16.7
23	2.5	19.2
24	2.6	21.8
25	2.9	24.7
26	3.0	27.6
27	3.0	30.6
28	3.1	33.7
29	3.2	36.9
30	3.5	40.4
31	3.7	44.1
32	3.7	47.8
33	3.5	51.3
34	3.5	54.8
35	3.6	58.4
36	3.8	62.2
37	3.9	66.1
38	3.7	69.8
39	3.9	73.7
40	3.7	77.5
41	3.7	81.2
42	3.9	85.1
43	3.4	88.5
44	3.2	91.7
45	2.8	94.5
46	2.4	96.9
47	1.6	98.6
48	1.0	99.5
49	0.4	99.9
50	0.1	100.0

Table 15 presents the percentage of students choosing each alternative (A, B, C, and D) for each multiple-choice question on the Grade 6 English Language Arts Part B: Reading (Multiple Choice). The correct response (key) for each question is also identified. The results presented in Table 15 can best be used in conjunction with results presented in jurisdiction and school reports to interpret strengths and weaknesses of local programs.

Table 15
Grade 6 English Language Arts
Results for Individual Multiple-Choice Questions

		Distribution of Responses* (%)						Distribution of Responses* (%)			
Item	Key	A	B	C	D	Item	Key	A	B	C	D
1	C	4.5	4.2	67.3	24.0	26	A	49.2	14.3	14.3	22.0
2	D	7.6	4.6	3.9	83.7	27	A	60.3	12.4	20.7	6.4
3	C	5.7	9.5	79.6	5.0	28	D	9.2	13.5	7.0	69.9
4	B	12.8	54.0	27.1	5.9	29	C	11.4	14.5	66.7	7.1
5	A	63.9	9.8	15.1	11.0	30	B	12.0	57.6	11.5	18.3
6	B	8.7	86.4	3.1	1.7	31	D	6.1	4.2	8.9	80.5
7	B	23.2	69.7	2.0	5.0	32	C	5.5	6.9	80.9	6.4
8	D	7.6	19.4	2.2	70.7	33	B	8.2	76.6	8.6	6.3
9	A	70.1	3.2	20.7	5.9	34	B	5.8	65.5	24.0	4.3
10	A	88.4	3.6	5.9	1.5	35	D	6.0	8.7	15.6	69.4
11	C	7.7	5.9	82.2	4.1	36	D	18.6	19.5	9.5	52.0
12	D	20.0	3.2	16.6	60.2	37	A	37.1	11.0	7.9	43.3
13	D	3.9	9.2	9.3	77.4	38	C	6.9	7.4	70.9	14.2
14	B	3.6	88.2	3.3	4.8	39	A	64.7	11.0	12.5	11.2
15	A	88.3	6.9	3.2	1.5	40	D	19.3	8.1	10.2	61.8
16	C	2.5	6.0	69.8	21.6	41	B	50.5	34.4	10.6	3.8
17	A	81.4	2.5	9.4	6.6	42	C	45.9	13.2	33.6	6.7
18	B	3.8	72.9	20.7	2.5	43	B	13.0	61.5	15.3	9.3
19	C	16.8	14.7	53.8	14.5	44	D	8.0	5.0	30.0	56.1
20	B	15.2	67.8	6.4	10.2	45	C	10.3	28.5	47.4	12.2
21	C	7.6	8.1	71.9	12.3	46	A	65.5	11.0	17.2	4.8
22	B	12.0	49.7	14.6	23.6	47	C	7.7	11.6	65.4	13.8
23	A	46.9	37.8	5.2	9.9	48	D	15.2	11.7	13.5	58.1
24	B	16.1	42.3	20.9	20.5	49	C	21.5	5.6	51.9	19.3
25	D	16.0	18.7	12.7	52.4	50	A	60.1	15.8	6.7	15.6

*The sum of the percentages for each question may be less than 100% because the No Response category is not included. The No Response category does not exceed 1.7% for any one of these questions.

Questions on the Grade 6 English Language Arts Part B: Reading (Multiple Choice) ranged from those that students found to be quite difficult, such as question 42 with only 33.6% of the students answering correctly, to those that were quite easy for students, such as question 10 with 88.4% of the students answering correctly. An example of a question of average difficulty is question 39 with 64.7% of the students answering correctly.

Table 16 indicates the percentage of jurisdictions classified as significantly below, not different from, or significantly above the provincial average (see page 5) for each subtest, for Part A: Composition, and for Part B: Reading (Multiple Choice), and for the total combined score.

Table 16
Grade 6 English Language Arts
Distribution of Jurisdiction Levels of Achievement

Subtest	Distribution of Jurisdiction Levels of Achievement: Below Provincial Average (%)	Not Different from Provincial Average (%)	Above Provincial Average (%)
Part A: Composition			
Content	6.1	83.0	10.9
Development	5.2	82.2	12.6
Sentence Structure	8.7	78.3	13.0
Vocabulary	7.4	80.9	11.7
Conventions	11.3	75.2	13.5
Part A: Total Score	14.3	75.2	10.4
Part B: Reading (Multiple Choice)			
Reporting Category			
Identifying the Organization of Ideas	15.2	75.2	9.6
Analysing Details	14.8	73.5	11.7
Associating Meanings	15.7	70.9	13.5
Synthesizing Ideas	11.7	76.5	11.7
Process Category			
Literal Understanding	15.7	74.8	9.6
Inferential Understanding	16.1	71.3	12.6
Evaluation	16.1	71.3	12.6
Part B: Total Score	16.5	68.7	14.8
Total Combined Score	15.2	70.0	14.8

SECTION 6:
GRADE 9 MATHEMATICS

THE GRADE 9 MATHEMATICS ACHIEVEMENT TEST

General Description

The Grade 9 Mathematics Achievement Test and the Grade 9 Mathematics Pilot Test were administered June 14, 1988. Information in this section is limited to the Grade 9 Mathematics Achievement Test.

The Grade 9 Mathematics Achievement Test consisted of 75 multiple-choice questions answered by students on machine-scorable answer sheets. The time allotted for writing the Grade 9 Mathematics Achievement Test was 90 minutes.

The scope of the Grade 9 Mathematics Achievement Test was limited to curriculum objectives that could be efficiently measured on a paper and pencil test.

Content of the Test

The Grade 9 Mathematics Achievement Test was designed to reflect the *Grade 9 Mathematics Curriculum Specifications* (revised May 1987). These specifications reflected the content objectives in the *Program of Studies for Junior High Mathematics, 1978*. All test questions were drawn from the content of the six strands in the core program prescribed for Grade 9. These six strands and the number of questions allotted to each strand were:

- Number Systems (17 questions)
- Ratios and Proportions (8 questions)
- Measurement (5 questions)
- Geometry (14 questions)
- Graphing (5 questions)
- Algebra (26 questions).

Questions on the Grade 9 Mathematics Achievement Test measured student achievement at four cognitive levels. These were:

- Knowledge (recall, memorize, perform routine computations)
- Comprehension (perform algorithms, demonstrate understanding of principles and concepts, interpret, translate)
- Application (formulate the problem symbolically, then manipulate the symbolic representation according to some previously learned algorithms)
- Higher Mental Activities (solve novel problems, prove new results, generate new algorithms).

Table 17 presents the blueprint used to develop the Grade 9 Mathematics Achievement Test. Classification by strand (reporting category) and cognitive level for each question included in the test is indicated in the table.

Table 17
Grade 9 Mathematics
Achievement Test Blueprint

Strand (Reporting Category)	Cognitive Level				Emphasis (%)
	Knowledge	Comprehension	Application	Higher Mental Activities	
Number Systems	*1,2,3,4, 11,12*	*6,7,8,9,10, 17*	*5,13,14, 15*	*16*	22.6
Ratios and Proportions	*18*	*19,20*	*21,23,24, 25*	*22*	10.7
Measurement	*26*		*27,28,29*	*30*	6.7
Geometry	*31,37,38*	*35,36,39,40*	*32,33,41, 42,43,44*	*34*	18.7
Graphing		*45*	*46,47,48, 49*		6.7
Algebra	*60,62,63, 65*	*50,51,56,57 61,64,66,67, 68,69,70,71, 72,74,75*	*52,54,55, 58,73*	*53,59*	34.6
Emphasis (%)	20.0	37.3	34.7	8.0	100.0

ABSENTEES AND EXEMPTIONS

Table 18 presents the number and the percentage distribution of students who wrote the Grade 9 Mathematics Achievement Test, as well as the number and the percentage distribution of students who were absent or who were exempted from writing for various reasons.

Table 18
Grade 9 Mathematics
Students Tested, Absentees, and Exemptions

	Category	Number of Students	Per Cent
Grade 9 Mathematics Achievement Test	Students Tested	25 290	83.9
	Students Absent	1 128	3.7
	Students Exempt:		
	• Special needs	831	2.8
	• Not taught subject this term	108	0.4
	• ESL classes	91	0.3
	• Other (Approved by Director)	38	0.1
	Total Exempt	1 068	3.6
Grade 9 French Translation	*Students Tested (Language of instruction not English)	773	2.6
Grade 9 Pilot Test	*Students Tested	1 867	6.2

Note: The data are incomplete as not all schools returned absentee and exemption reports.
*The results for these students are not included in the provincial report.

RESULTS

Information in all tables contained in this section is based on the results achieved by 25 290 students. The students who wrote the Grade 9 Mathematics Achievement Test in 1988 were judged to be representative of the total population. Since over 24 900 of these students completed all questions on the test, it was concluded that sufficient time was allotted for writing the test.

The minimum acceptable and outstanding performance levels of achievement were established by a standard-setting procedure (see Appendix A). For the Grade 9 Mathematics Achievement Test, the standards established are such that to meet

- the minimum acceptable performance level, students must achieve a raw score of 35 out of 75
- the outstanding performance level, students must achieve a raw score of 63 out of 75.

Based on these standards, the results revealed that

- 88.9 per cent of students performed at or above the minimum acceptable performance level
- 24.3 per cent of students performed at or above the outstanding performance level.

These levels of performance were higher than were expected prior to administration of the test.

Provincial scores, in percentage points, are as follows:

- Provincial Average - 69.3
- Standard Deviation - 17.2

The KR-20 reliability coefficient for the Grade 9 Mathematics Achievement Test was 0.92.

As outlined in the blueprint presented on page 30, the questions on the Grade 9 Mathematics Achievement Test have been grouped according to subtests, strands (or reporting categories), and cognitive levels.

Provincial averages for these reporting categories and for the total test are presented in Table 19. Provincial averages were computed and rounded to one decimal. Consequently, the sum of the averages for the subtests is not exactly the same as the average for the total test.

Table 19
Grade 9 Mathematics
Raw Score Results by Subtest

Subtest	Number of Questions	Average	Standard Deviation
Total Test	75	51.9	12.9
Strand (Reporting Category)			
Number Systems	17	12.3	3.2
Ratios and Proportions	8	5.5	1.8
Measurement	5	3.3	1.2
Geometry	14	9.9	2.9
Graphing	5	3.3	1.2
Algebra	26	17.7	5.5
Cognitive Level			
Knowledge	15	11.0	2.7
Comprehension	28	19.8	5.5
Application	26	17.5	5.1
Higher Mental Activities	6	3.6	1.4

The percentage of students who obtained each score on the Grade 9 Mathematics Achievement Test (relative frequency) and the percentage of students scoring at or below each score (cumulative frequency) are presented in Table 20.

Table 20
Grade 9 Mathematics
Frequency Distribution of Raw Scores

Raw Score	Relative Frequency (%)	Cumulative Frequency (%)	Raw Score	Relative Frequency (%)	Cumulative Frequency (%)
0	0.0	0.0	38	1.7	17.1
1	0.0	0.0	39	1.8	18.9
2	0.0	0.0	40	1.7	20.6
3	0.0	0.0	41	1.9	22.4
4	0.0	0.0	42	2.0	24.4
5	0.0	0.0	43	1.9	26.3
6	0.0	0.0	44	2.2	28.5
7	0.0	0.0	45	2.2	30.7
8	0.0	0.0	46	2.5	33.2
9	0.0	0.0	47	2.2	35.4
10	0.0	0.0	48	2.4	37.8
11	0.0	0.0	49	2.2	40.1
12	0.0	0.1	50	2.6	42.6
13	0.0	0.1	51	2.5	45.1
14	0.0	0.1	52	2.5	47.6
15	0.1	0.2	53	2.7	50.3
16	0.1	0.2	54	2.6	52.9
17	0.1	0.3	55	2.8	55.7
18	0.1	0.5	56	3.0	58.7
19	0.2	0.7	57	2.7	61.4
20	0.2	0.8	58	2.7	64.1
21	0.3	1.1	59	3.0	67.1
22	0.3	1.4	60	2.9	69.9
23	0.4	1.8	61	3.0	73.0
24	0.5	2.3	62	2.7	75.7
25	0.5	2.8	63	2.8	78.5
26	0.6	3.5	64	2.6	81.0
27	0.8	4.3	65	2.5	83.5
28	0.7	4.9	66	2.6	86.1
29	0.8	5.7	67	2.3	88.4
30	0.9	6.6	68	2.4	90.9
31	1.1	7.7	69	2.3	93.1
32	1.0	8.7	70	2.0	95.1
33	1.2	9.9	71	1.6	96.7
34	1.2	11.1	72	1.5	98.1
35	1.3	12.4	73	0.9	99.1
36	1.4	13.8	74	0.6	99.7
37	1.5	15.4	75	0.3	100.0

Table 21 presents the percentage of students choosing each alternative (A, B, C, and D) for each multiple-choice question on the Grade 9 Mathematics Achievement Test. The correct response (key) for each question is also identified.

Table 21
Grade 9 Mathematics
Results for Individual Multiple-Choice Questions

Item	Key	Distribution of Responses* (%)				Item	Key	Distribution of Responses* (%)			
		A	B	C	D			A	B	C	D
1	C	0.2	4.9	93.0	1.9	39	A	74.6	7.0	11.4	6.8
2	C	16.8	8.8	71.1	3.2	40	C	10.6	6.0	52.9	30.5
3	A	64.8	15.4	10.0	9.7	41	B	4.6	68.2	11.2	15.9
4	A	40.1	21.8	3.9	34.0	42	D	5.1	12.0	13.7	69.1
5	A	59.3	2.8	30.2	7.5	43	A	59.3	18.0	14.6	7.8
6	D	3.9	2.9	7.2	86.0	44	B	20.6	60.8	14.4	3.8
7	D	6.3	31.6	9.4	52.2	45	D	10.9	4.5	5.3	79.1
8	D	18.9	2.4	7.9	70.8	46	A	78.4	7.2	9.3	4.9
9	B	20.0	55.6	16.6	7.3	47	B	18.9	42.5	19.6	18.5
10	D	16.3	4.3	9.2	70.0	48	D	8.2	13.1	31.7	46.7
11	B	3.3	78.5	5.0	13.1	49	C	7.2	8.0	78.4	6.1
12	C	8.9	4.8	82.2	4.0	50	D	8.5	13.7	17.5	59.9
13	B	13.2	72.9	8.7	4.8	51	A	73.5	15.3	6.6	4.3
14	A	74.4	11.0	9.5	4.8	52	A	61.3	24.0	11.4	3.0
15	A	91.7	2.5	3.8	1.9	53	B	24.2	36.2	31.7	7.0
16	C	1.5	6.4	84.2	7.6	54	C	8.4	16.7	70.1	4.5
17	D	8.2	5.4	2.2	84.2	55	B	13.6	71.8	12.1	2.1
18	C	33.1	3.3	59.7	3.9	56	A	68.6	9.4	18.3	3.4
19	B	1.3	94.2	3.8	0.7	57	C	7.5	28.9	59.2	4.0
20	B	13.6	68.5	4.6	13.1	58	B	5.8	60.4	21.4	11.9
21	C	23.2	12.0	58.6	6.2	59	B	12.4	70.5	10.6	5.6
22	C	4.3	29.9	60.3	5.0	60	C	4.1	12.6	77.9	5.0
23	B	30.3	56.0	12.3	1.2	61	C	7.8	9.7	68.0	14.1
24	B	5.9	73.3	10.7	9.8	62	D	2.5	4.5	13.6	79.0
25	B	5.2	82.1	7.4	5.0	63	C	10.0	5.5	75.7	8.3
26	C	5.1	18.5	70.9	5.4	64	A	88.5	2.8	3.5	4.6
27	C	33.8	3.0	60.9	2.3	65	D	6.6	16.8	7.1	69.0
28	D	4.1	16.6	6.5	72.5	66	D	3.3	5.5	18.6	72.0
29	A	79.1	10.4	3.8	6.6	67	A	74.0	11.1	9.6	4.6
30	B	11.0	43.8	38.2	6.1	68	D	3.4	7.1	18.0	71.0
31	D	11.9	5.1	15.3	67.6	69	D	10.0	9.6	6.4	73.3
32	D	2.5	4.1	24.3	69.0	70	B	11.9	66.0	12.6	8.9
33	C	3.1	12.0	71.0	13.4	71	C	7.6	9.9	66.4	15.4
34	A	68.2	17.5	9.0	4.8	72	D	5.3	7.6	11.9	74.4
35	C	4.7	17.0	74.7	3.4	73	A	58.7	7.1	23.5	10.0
36	B	8.0	85.5	4.9	1.5	74	A	70.6	8.0	6.9	13.7
37	C	7.2	2.4	81.6	8.6	75	A	49.8	22.1	18.1	9.0
38	D	1.9	6.8	2.0	89.2						

*The sum of the percentages for each question may be less than 100% because the No Response category is not included. The No Response category does not exceed 1.0% for any one of these questions.

Questions on the Grade 9 Mathematics Achievement Test ranged from those that students found to be quite difficult, such as question 53 with only 36.2% of the students answering correctly, to those that were quite easy for students, such as question 19 with 94.2% of the students answering correctly. An example of a question of average difficulty is question 42 with 69.1% of the students answering correctly.

Table 22 indicates the percentage of jurisdictions classified as significantly below, not different from, or significantly above the provincial average (see page 5) for each subtest and for the total test.

Table 22
Grade 9 Mathematics
Distribution of Jurisdiction Levels of Achievement

	Distribution of Jurisdiction Levels of Achievement		
Subtest	Below Provincial Average (%)	Not Different from Provincial Average (%)	Above Provincial Average (%)
Total Test	24.9	60.1	15.0
Strand (Reporting Category)			
Number Systems	21.1	59.6	19.2
Ratios and Proportions	17.4	72.3	10.3
Measurement	19.2	71.4	9.4
Geometry	21.1	60.6	18.3
Graphing	23.9	62.4	13.6
Algebra	24.4	61.5	14.1
Cognitive Level			
Knowledge	26.3	57.7	16.0
Comprehension	24.9	58.2	16.9
Application	21.1	64.3	14.6
Higher Mental Activities	7.5	84.0	8.5

SECTION 7: SUMMARY OF RESULTS

In Table 23, the percentage of students performing at or above the minimum acceptable performance level and the outstanding performance level are summarized, as are the provincial average and standard deviation for the June 1988 achievement tests (Grade 3 Social Studies, Grade 6 English Language Arts, and Grade 9 Mathematics).

Table 23
Summary of Provincial Averages, Standard Deviations, and Performance
June 1988 Achievement Tests

Achievement Test	Percentage of students at (or above) Minimum Acceptable Performance Level	Percentage of students at (or above) Outstanding Performance Level	Provincial Average (%)	Standard Deviation (%)
Grade 3 Social Studies	80.0	18.8	67.6	18.2
Grade 6 English Language Arts	83.1	9.9	62.5	15.2
Grade 9 Mathematics	88.9	24.3	69.3	17.2

It should be noted that the standards in Grade 6 English Language Arts for the total test were created by combining the results of the standard setting for Part B: Reading (reported in Section 5) with the standards implicit in the scoring system for Part A: Composition.

Students whose writing skills were on the borderline of adequacy were equally likely to receive a score of Satisfactory (3) or Limited (2) on each written-response scale, for an average of 2.5 on each scale, or 12.5 over five scales. Since there were 25 possible marks, contributing 50% of the total combined score, borderline adequate students received 25 marks for written response in their total combined score. Adding this to the standard of 23 marks for the multiple choice set the standard at 48% for the total combined score.

For the outstanding level of achievement, a student received an average mark of Proficient (4), for a total of 20 over five scales. Since there were 25 possible marks, contributing 50% of the total combined score, borderline outstanding students received 40 marks for written response in their total combined score. Adding this to the standard of 43 marks for the multiple choice set the standard at 83% for the total combined score.

Schools and jurisdictions wishing to determine the percentage of their students achieving these standards on the total combined score may consult the table of frequency distributions for the total combined score included in the supplementary school and jurisdiction reports for Grade 6 English Language Arts. The cumulative frequency column in this table shows the percentage of students in the school or jurisdiction scoring at or below a particular percentage. By subtracting the percentage of students scoring 47% or less from 100%, the percentage achieving the standard for borderline adequate achievement can be calculated. By subtracting the percentage of students scoring 82% or less from 100%, the percentage achieving the standard for borderline outstanding achievement can be calculated.

SECTION 8:
STUDIES OF CHANGES IN STUDENT ACHIEVEMENT OVER TIME

INTRODUCTION

Purpose

An important goal of Alberta Education is to measure and report changes in student achievement. Comparing student performance on the 1984 and the 1988 achievement tests is one way of meeting this goal. A direct comparison of the average scores on the 1984 and the 1988 tests cannot be made; although the achievement tests were parallel in form and content, their levels of difficulty may not have been equal. A study was therefore undertaken to permit comparisons of student performance in 1984 and 1988.

Study Design

The study consisted of two parts for the Grade 3 Social Studies and the Grade 6 English Language Arts tests: test equating and standard setting. The Grade 9 study consisted of three parts: test equating, standard setting, and comparing equivalent items.

To equate the 1984 and the 1988 test scores, samples of grades 3, 6, and 9 students, who also completed the 1988 achievement tests, wrote the 1984 Grade 3 Social Studies, Grade 6 English Language Arts, and Grade 9 Mathematics Achievement tests respectively. The students selected for the study were from schools chosen to be representative of rural, urban, large, and small schools throughout the province.

The 1984 and the 1988 multiple-choice scores from these students were used to establish what the 1984 equivalent scores would be on the 1988 test. (For an explanation of the test-equating process, refer to Appendix B of this report.) Based on this equating process, the multiple-choice scores for the 1988 population were converted to estimated equivalent 1984 scores, and the 1984 and the 1988 populations were then compared.

A different procedure was used to equate the written-response section of the 1984 and the 1988 English Language Arts Achievement tests. The 1984 papers, written by the 1988 sample of students, were scored by the same markers who scored the 1988 tests, using the 1988 scoring descriptors. A sample of papers from the 1984 population was rescored at the same time. Two analyses were done on the data from these marking sessions, one which compared assignment difficulty and another which compared marking standards. The results of these analyses were used to adjust the frequency distributions of scores for the 1988 population to make the distributions comparable to the scores for the 1984 population. These adjustments served to correct for differences in test difficulty and scoring standards.

The standard-setting part of the study involved groups of experienced teachers of grades 3, 6, and 9. Their task was to establish the scores that represent the minimum acceptable and the outstanding levels of performance on each of the achievement tests in 1984 and in 1988. The percentage of students meeting each of the standards set for the respective tests was then compared.

In addition to test equating and standard setting, a third procedure was used to compare student achievement in Grade 9 Mathematics. This procedure involved comparing the achievement on equivalent items in 1984 with those of 1988. Fifteen items from the 1984 test were paired with 15 equivalent items from the 1988 test. The average score obtained by the 1984 and the 1988 populations of students on their respective sets of equivalent items was compared.

Limitations of the Study

Many factors other than changes in levels of achievement could have produced differences in student performance on the 1984 and the 1988 achievement tests. Included were factors such as practice effects, different use by teachers of the bulletins provided by the Student Evaluation and Records Branch, and changes in examinee motivation. Because the effects of these extraneous variables could not be controlled, caution is required when drawing conclusions about the results of this study.

GRADE 3 SOCIAL STUDIES

Grade 3 Social Studies: Multiple Choice

The 1984 and the 1988 Grade 3 Social Studies Achievement tests each consisted of 50 multiple-choice questions in one machine-scorable booklet.

The average score in 1984 was 31.4 out of 50 possible marks; in 1988, the average score was 33.8 correct out of 50.

The special comparison study involved administering the 1984 test to a sample of students who would also write the 1988 test and using a single group of judges to determine the standards for both tests.

Comparison Through Re-Administering the 1984 Test

Sixteen schools in 14 jurisdictions were selected for the study. These schools were chosen to be representative of rural, urban, large, and small schools throughout the province. In June 1988, 296 Grade 3 students, who also completed the 1988 test, wrote the 1984 Social Studies Achievement Test. The 1984 and the 1988 multiple-choice scores from these students were used to determine what the equivalent scores on the 1984 test would be. (For an explanation of the test-equating process, refer to Appendix B of this report.) All multiple-choice scores for the students who wrote tests in 1988 were converted to estimated equivalent 1984 scores, and the 1984 Grade 3 students and the 1988 Grade 3 students were then compared. Table 24 on the next page shows the results of the comparison.

Table 24
Grade 3 Social Studies
Multiple-Choice Comparison of the 1984 and the 1988 Test Results

	1984 Test	1988 Test
Number of Questions	50	50
Number of Students Writing	25 775	29 961
Average Score after Equating	31.5	31.1
Standard Deviation after Equating	8.4	9.1

The average 1988 score was 0.4 out of 50 lower than the average 1984 score. A two-tailed t-test showed that this difference was statistically significant beyond the 0.001 level of probability. However, the reliability of the estimated scores is limited. Considering this limitation, the difference may not reflect a real difference in achievement scores.

Standard Setting

Twenty-one experienced Grade 3 Social Studies teachers from schools throughout the province were selected to participate in the standard-setting section of the study. Please refer to Appendix A for an explanation of the standard-setting procedure.

These teachers reviewed both the 1984 and the 1988 tests. The teachers set two standards for each test: the minimum acceptable level of achievement and the outstanding level of achievement. Table 25 shows the results of the standard-setting process.

Table 25
Grade 3 Social Studies
Multiple-Choice Standard-Setting Results

		Minimum Acceptable Level of Achievement		Outstanding Level of Achievement	
Test	Maximum Score	Raw Score Representing Standard	Per Cent of Students Meeting Standard	Raw Score Representing Standard	Per Cent of Students Meeting Standard
1984	50	25	77.9	40	18.2
1988	50	26	80.0	43	18.8

The results revealed that nearly the same number of students met the two standards in 1984 and in 1988. Since a change of one in the raw score representing the standard will result in a change of approximately four per cent in the percentage of students meeting the standard, the differences shown are not important.

The standard-setting results showed that, in the judgment of the standard setters, students achieved the minimum acceptable level and the outstanding level of achievement as frequently in both test administrations.

Conclusion

The results of the test-equating analysis indicated that the average achievement may have been lower in 1988 than in 1984; however, standard-setting results indicated that similar percentages of students achieved an adequate or outstanding level. Given the limitations on the precision of the test-equating process with an equating group of 296 students, there is insufficient evidence to indicate that there is a change in achievement between 1984 and 1988. The conclusion, therefore, is that the level of achievement is essentially the same as it was in 1984.

GRADE 6 ENGLISH LANGUAGE ARTS

Grade 6 English Language Arts: Multiple Choice

The 1984 and the 1988 Grade 6 English Language Arts Achievement tests consisted of two parts, Part A: Composition and Part B: Reading (Multiple Choice). On both tests, Part B: Reading contained 50 multiple-choice questions.

The average score in 1984 was 33.5 out of 50 possible marks; in 1988, the average score was 32.4 correct out of 50.

The special comparison study involved administering the 1984 test to a sample of students who would also write the 1988 test and using a single group of judges to determine the standards for both tests.

Comparison through Re-administering the 1984 Test

In June 1988, 205 Grade 6 students from eight schools in six jurisdictions were selected for test-equating purposes. (For an explanation of the test-equating process, refer to Appendix B.) These schools were chosen to be representative of rural, urban, large, and small schools throughout the province. Based on this test-equating process, the multiple-choice scores for the 1988 population were converted to estimated equivalent 1984 scores, and the 1984 and the 1988 populations were then compared. Table 26 on the next page shows the results of the comparison.

Table 26
Grade 6 English Language Arts
Multiple-Choice Comparison of the 1984 and the 1988 Test Results

	1984 Test	1988 Test
Number of Questions	50	50
Number of Students Writing	28 812	32 069
Average Score after Equating	33.5	34.7
Standard Deviation after Equating	8.1	8.0

The average 1988 score was 1.2 out of 50 higher than the average 1984 score. A two-tailed t-test showed that this difference was statistically significant beyond the 0.001 level of probability.

Standard Setting

Twenty-one experienced Grade 6 English Language Arts teachers from schools throughout the province were selected to participate in the standard-setting section of the study. These teachers were among those marking the written-response part of the test. Please refer to Appendix A for an explanation of the standard-setting procedure.

These teachers reviewed the multiple-choice parts of both the 1984 and the 1988 tests. The teachers set two standards for each test: the minimum acceptable level of achievement and the outstanding level of achievement. Table 27 shows the results of the standard-setting process.

Table 27
Grade 6 English Language Arts
Multiple-Choice Standard-Setting Results

		Minimum Acceptable Level of Achievement		Outstanding Level of Achievement	
Test	Maximum Score	Raw Score Representing Standard	Per Cent of Students Meeting Standard	Raw Score Representing Standard	Per Cent of Students Meeting Standard
1984	50	25	84.8	42	15.8
1988	50	23	83.3	43	14.9

The results revealed that nearly the same number of students met the two standards in 1984 and in 1988. Since a change of one in the raw score representing the standard will result in a change of approximately four per cent in the percentage of students meeting the standards, the differences shown are not important.

The standard-setting results showed that, in the judgment of the standard setters, students achieved the minimum acceptable level and the outstanding level of achievement as frequently in both test administrations.

Conclusion

The results of the test-equating analysis indicated that achievement on the reading section on average was higher in 1988 than in 1984, but similar percentages of students achieved an adequate or outstanding level. It was concluded that achievement on reading in 1988 was higher than in 1984.

Grade 6 English Language Arts: Written Response

Equating the written-response sections of the 1984 and the 1988 English Language Arts Achievement tests required different procedures from those used for the multiple-choice comparisons. See Appendix C for a detailed description of procedures and analyses.

Results

The analyses indicated that the assignment was more challenging in 1988, and that the marking standards were higher.

Data from analyses were used to adjust the frequency distributions of scores for the 1988 population to make them comparable to the scores for the 1984 population.

The percentage of students scoring Satisfactory (3) or higher on each scale in 1988 was then compared to the adjusted percentage of students scoring Satisfactory (3) or higher on each scale in 1984. Table 28 below shows the results of this comparison.

Table 28
Grade 6 English Language Arts: Written-Response
Frequencies of Scores of Satisfactory or Higher in Per Cent

	Content	Development	Sentence Structure	Vocabulary	Convention
1984 Results	79.6	78.4	82.3	81.3	78.2
1988 Results (Adjusted)	90.4	75.4	86.6	85.8	76.6
Difference	10.8	-3.0	4.4	4.5	-1.6

Based on 1984 standards, the results indicate that performance in 1988 was higher in Content, Sentence Structure, and Vocabulary but was lower in Development and Convention.

Limitations

The differences were not subjected to tests of statistical significance as there were no standard tests for this type of analysis. Because of the limited size of the groups used to calculate score adjustments, the adjustments that were made were necessarily limited in precision. A change of one student in the number attaining a score results in a change of half of one per cent in the adjustments. Single-marker scoring of written response, as used in these tests, is known to be less reliable than multiple-marker scoring (as used in the Diploma Examinations in language arts, for example) or machine scoring of multiple-choice questions.

Conclusions

Given the limitations of the study, it is reasonable to conclude that written-response achievement has increased slightly in 1988 compared to 1984. Based on the test-equating process (see Appendix C), the one area that increased significantly is Content, where adjusted 1988 scores were markedly higher than the 1984 scores, although actual scores were lower in 1988. Two factors have been identified which would help to explain the lower actual written-response scores in 1988: marking standards in 1988 were much higher than in 1984, and the assignment was more difficult in 1988. Teacher comments during the 1988 marking session support this explanation of lower 1988 scores. In fact, teacher comments during the marking session indicate that the overall quality of student writing is at least as good and may be of higher quality in 1988 than in 1984.

GRADE 9 MATHEMATICS

Grade 9 Mathematics: Multiple Choice

In 1988, the Grade 9 Mathematics Achievement Test was administered to 25 290 students enrolled in the current Grade 9 mathematics program. A total of 1 867 students were exempted from the achievement test because they were enrolled in the new Grade 9 mathematics program.

An analysis was carried out to determine whether students from schools where the 1988 achievement test was written were representative of the 1988 Grade 9 population. Results revealed that these students were representative of the Grade 9 population.

The Grade 9 Mathematics Achievement tests in 1984 and in 1988 consisted of 75 questions*. The average score in 1984 was 66.3%; in 1988, the average score was 69.3%. These scores were based on the results of 27 121 students in 1984 and 25 290 students in 1988.

The special comparison study involved administering the 1984 test to a sample of students who would also write the 1988 test, using a single group of judges to determine the standards for both tests, and comparing equivalent questions.

Comparison through Re-administering the 1984 Test

Seven schools in seven jurisdictions were selected to participate in the study. These schools were chosen to be representative of rural, urban, large, and small schools throughout the province. In June 1988, 289 Grade 9 students, who also completed the 1988 test, wrote the 1984 Grade 9

---

*Two questions were deleted from the 1984 achievement test.

Mathematics Achievement Test. The 1984 and the 1988 multiple-choice scores from these students were used to determine what the equivalent scores on the 1988 test would be on the 1984 test. (For an explanation of the test-equating process, refer to Appendix B of this report.) All multiple-choice scores for the 1988 population were converted to estimated equivalent 1984 scores, and the 1984 and the 1988 populations were then compared. Table 29 shows the results of the comparison.

Table 29
Grade 9 Mathematics
Multiple-Choice Comparison of the 1984 and the 1988 Test Results

	1984 Test	1988 Test
Number of Questions	73	75
Number of Students Writing	27 121	25 290
Average Score after Equating	66.3%	67.7%
Standard Deviation after Equating	17.8%	17.4%

The average 1988 score was 1.4 per cent higher than the average 1984 score. A two-tailed t-test showed that this difference was statistically significant beyond the 0.001 level of probability.

Standard Setting

Twenty-one experienced Grade 9 Mathematics teachers from schools throughout the province were selected to participate in the standard-setting section of the study. Please refer to Appendix A for an explanation of the standard-setting procedure.

These 21 teachers reviewed the multiple-choice parts of both the 1984 and the 1988 tests. The teachers set two standards for each test: the minimum acceptable level of achievement and the outstanding level of achievement. The following table shows the results of the standard-setting process.

Table 30
Grade 9 Mathematics
Multiple-Choice Standard-Setting Results

		Minimum Acceptable Level of Achievement		Outstanding Level of Achievement	
Test	Maximum Score	Raw Score Representing Standard	Per Cent of Students Meeting Standard	Raw Score Representing Standard	Per Cent of Students Meeting Standard
1984	73	33	86.2	62	18.7
1988	75	35	88.9	63	24.3

The results revealed that a slightly larger percentage of students scored at or above the minimum acceptable level of achievement in 1988 than students did in 1984. A considerably larger percentage of students scored at or above the minimum outstanding level of achievement in 1988 than students did in 1984.

Equivalent Questions

Fifteen questions from the 1984 achievement test were judged to be equivalent to 15 questions in the 1988 achievement test by the Grade 9 Mathematics Test Review Committee and by the 21 teachers who participated in the standard-setting process.

Questions are judged to be equivalent if they have all of the following characteristics:

1. They test the same specific curriculum objectives.
2. The questions require the same level of reading skills.
3. The wordings of the questions are very similar in terms of their complexity and the amount of information they present.
4. The questions require the same number of steps or thought processes.
5. All three wrong alternatives are similar in terms of the processes required to arrive at those answers.

The average score on the 15 questions was 70.5% in 1984 and 73.7% in 1988.

Conclusion

The results obtained from the methods used to analyse and compare the 1984 and the 1988 test results indicated that on the average, students achieved a higher level of performance in 1988 than in 1984.

APPENDIX A: STANDARD-SETTING PROCEDURES

Rationale

The purpose of standard setting in the Achievement Testing Program is to answer the question of whether province-wide performance is satisfactory. To use standard setting in this way requires two distinct judgments. The first is to establish what percentage of the students tested can be expected to achieve at least the minimum level of skill and knowledge required to proceed to the next level in that subject, assuming there are adequate teaching and resources. The second is to establish the test score that represents that minimum level.

Satisfactory provincial performance can be said to occur when the percentage of students scoring at or above the established test score is equal to or greater than the expected percentage. Two similar judgments must be made for any other standards required, such as the outstanding level of skill and knowledge.

Standard setters must have a shared concept of the skills and knowledge of the borderline students for each standard set. Experience has shown that, in the absence of any other authoritative judgment, it is useful to estimate the expected percentage of students who will achieve the minimum level as 85%, and the expected percentage who will achieve the outstanding level as 15%.

Essential Elements of Procedures

1. Select standard setters who are familiar with both the curriculum and the characteristics of students who have written the test.

2. Explain the rationale for and the purpose of standard setting.

3. Reach a consensus on the characteristics of borderline students for each standard set, with emphasis on the characteristics that affect responses to the achievement test in question.

4. Have standard setters make and record question-by-question judgments in connection with the minimum skills standard.

5. Have standard setters make and record global judgments about the test in connection with the minimum skills standard and the outstanding skills standard.

6. Calculate the raw score derived from question-by-question judgments for each standard setter. Inform the standard setters of their individual standards and of the median for the group of standard setters. Also inform them of the median result of their global judgments.

7. Present the standard setters with data on the actual distribution of scores and the actual response frequencies.

8. Allow (but do not force) the standard setters to revise their judgments, stressing that they need not consider the actual results.

APPENDIX B:
EQUATING MULTIPLE-CHOICE TEST SCORES

Comparing achievement in two different groups requires some common measure. The Student Evaluation and Records Branch develops new achievement tests for each administration, reflecting changes in curricular emphases and refinements in test design learned from earlier tests. Thus scores from one administration of an achievement test cannot be directly compared to scores from another administration.

Various techniques are available to address this problem. In order to compare the 1988 achievement test results with results for the same subjects in 1984, the branch chose a variation of test equating as one of the techniques to be used.

Each 1984 test was administered in 1988 to a sample of students who were also writing the 1988 achievement test in that subject. The 1984 tests were administered either one week before or one week after the 1988 tests were written, using the same instructions that were used in 1984. These students had not been exposed to the 1984 test prior to administration.

The 1984 tests that had been re-administered were scored using the same keys that were used in 1984. Scores were matched by student name and school with their scores on the 1988 achievement tests. Students for whom only one score was available were removed from the sample.

The two sets of scores in the sample were then assumed to represent the same range of achievement, as the same sample of students had produced both sets; both tests measured achievement in the same curriculum. Thus, achievement at a particular level in one set of scores should be equivalent to achievement at the same level in another set of scores. A score at the 20th percentile on the 1988 test would be equivalent to a score at the 20th percentile on the 1984 test, because the percentiles were based on exactly the same students. (A percentile score represents the percentage of the sample or population scoring at or below that particular score.)

For example, a score of 31 on the 1984 English Language Arts Achievement Test fell at the 23.9th percentile of the sample group. On the 1988 test, the 23.9th percentile came between a score of 26 (percentile ranking of 21.7) and 27 (percentile ranking of 25.1) in the sample group. Thus the score of 31 on the 1984 test was assigned an equivalent score of 26.6 on the 1988 test.

This fact was used to calculate equivalent scores for the two tests. The 1984 test was used as the anchor test. Each 1984 score was converted to its percentile equivalent. Percentile scores for the 1988 scores in the sample were also calculated. A transformation table was then produced.

The transformation table was used to translate all scores in the 1988 population to their equivalents in the 1984 scores. It was then possible to treat the estimated scores as equivalent to the actual scores of the 1984 population and perform a comparison using the two-tailed t-test to determine whether there were statistically significant differences between the achievement scores of students who wrote the 1984 test in 1984 and students who wrote the 1988 test in 1988.

APPENDIX C: GRADE 6 ENGLISH LANGUAGE ARTS: WRITTEN-RESPONSE EQUATING

The written-response scores for the Grade 6 English Language Arts Achievement Test are reported on five scales, with possible scores ranging from 0 to 5 on each scale. These ranges are too limited for practical application of the within-group equipercentile technique that was used with the multiple choice (see Appendix B). Another consideration is the possible difference in the application of the scoring system between the 1984 scoring and the 1988 scoring. Scale descriptors have been revised, and marker training procedures have been further refined, so that scores cannot be assumed to be comparable. It was necessary to develop procedures to overcome these constraints.

Procedures

The 1984 written-response test was administered to a sample of 180 students who also wrote the 1988 written-response test. The sample was essentially the same as that used for the multiple-choice test equating, although some students completed only one section of the test. The 1984 tests were then scored using the 1988 scale descriptors by markers who had participated in the scoring of the 1988 written response. A sample of 295 papers written in 1984 were rescored at the same time by the same markers. Markers did not know the year in which any particular paper had been written.

Analyses

Results of the 1988 scoring of the 1984 papers written in 1988 were compared with results for the 1984 papers written in 1984. No significant differences in mark distributions were found, supporting the assumption that the scoring was unbiased.

Scores for 1988 students on the 1984 papers were compared to the same students' scores on the 1988 paper. This analysis indicated that the 1988 assignment was more difficult than the 1984 assignment; higher scores were more frequent on the 1984 assignment.

In 1988, a sample of 1984 student papers were rescored along with the 1984 papers that were administered in 1988. These scores were then compared. This analysis indicated that the 1988 standards were higher than the 1984 standards; higher scores were more frequent in the 1984 marking, and lower scores were more frequent in the 1988 marking.

The data from the two analyses, comparing assignment difficulty and marking standards, were used to adjust the frequency distributions of scores for the 1988 population to make them comparable to the scores for the 1984 population. For example, the difference between the percentage of Excellent (5) scores awarded for Content on the 1984 assignment and the percentage of Excellent (5) scores awarded for Content on the 1988 assignment in the group of students who wrote both was added to the percentage of Excellent (5) scores awarded for Content in the 1988 population. This adjustment was repeated for all scales and scores and served to correct for test difficulty. A similar procedure was used to adjust for the change in standards.

The percentage of students scoring Satisfactory (3) or higher on each scale in 1984 was then compared to the adjusted percentage of students scoring Satisfactory (3) or higher on each scale in 1988. Table 28 in Section 8 shows the results of this comparison.

ACHIEVEMENT TESTING PROGRAM
PROVINCIAL REPORT
QUESTIONNAIRE

It is the intention of the Student Evaluation and Records Branch to produce documents that will be useful to teachers and administrators. The purpose of the following questionnaire is to garner your opinions about the *Provincial Report* so that these opinions can be considered when the content and format of the report are reviewed prior to June 1989.

Please take a moment to respond to the following questions, then detach this sheet and send it to

Mr. Michael Robinson
Assistant Director, Data Analysis and
Student Records Services
Student Evaluation and Records Branch
Alberta Education
11160 Jasper Avenue
Edmonton, Alberta T5K 0L2

USE OF THE REPORT

1. (Please check the box beside the statement that applies to you.)

My present role is primarily that of

- teacher ☐ 1
- school administrator ☐ 2
- central office administrator ☐ 3
- school board member ☐ 4
- other (please specify)______________________________

2. (Please check the box beside the statement that applies to you.)

I read the report, but I DID NOT use it to interpret the results attained by students. ☐ 5

I read the report, and I used it to interpret the results attained by the students in

- my classroom ☐ 6
- my school ☐ 7
- my jurisdiction ☐ 8

3. (Please respond to the following statement if you have checked boxes 6, 7, or 8 above.)

I have made use of the results to alter the educational program offered in

- my classroom ☐ 9
- my school ☐ 10
- my jurisdiction ☐ 11

CONTENT OF THE REPORT

1. (Please judge the usefulness of the information included in the various sections of the report by checking the appropriate boxes below.)

	Very Useful	Adequate for Use	Of Some Use	Of No Use
Section 2: Student Achievement Testing Program	☐ 12	☐ 13	☐ 14	☐ 15
Section 3: Guidelines for Interpreting the Results of Achievement Tests	☐ 16	☐ 17	☐ 18	☐ 19
Sections 4-6: Specific Achievement Test Results	☐ 20	☐ 21	☐ 22	☐ 23
Section 7: Summary of Results	☐ 24	☐ 25	☐ 26	☐ 27
Section 8: Studies of Changes in Student Achievement Over Time	☐ 28	☐ 29	☐ 30	☐ 31

2. (If you wish, please comment further on the content of the report in the space below.)

FORMAT OF THE REPORT

1. (Please judge the usefulness of the report's format by checking the appropriate boxes below.)

	Very Useful	Adequate for Use	Of Some Use	Of No Use
Organization into Separate Sections	☐ 32	☐ 33	☐ 34	☐ 35
Double-Column Presentation of Text	☐ 36	☐ 37	☐ 38	☐ 39
Presentation of Tables	☐ 40	☐ 41	☐ 42	☐ 43
Blending of Information through Text and Tables	☐ 44	☐ 45	☐ 46	☐ 47

2. (If you wish, please comment further on the format of the report in the space below.)